ELLIPTIC LAW FOR REAL RANDOM MATRICES

ALEXEY NAUMOV[1]

ABSTRACT. In this paper we consider ensemble of random matrices $\mathbf{X}_n$ with independent identically distributed vectors $(X_{ij}, X_{ji})_{i \neq j}$ of entries. Under assumption of finite fourth moment of matrix entries it is proved that empirical spectral distribution of eigenvalues converges in probability to a uniform distribution on the ellipse. The axis of the ellipse are determined by correlation between X_{12} and X_{21}. This result is called Elliptic Law. Limit distribution doesn't depend on distribution of matrix elements and the result in this sence is universal.

CONTENTS

1. INTRODUCTION

Let us consider real random matrix $\mathbf{X}_n(\omega) = \{X_{ij}(\omega)\}_{i,j=1}^n$ and assume that the following conditions (**C0**) hold

a) Pairs $(X_{ij}, X_{ji}), i \neq j$ are independent identically distributed (i.i.d.) random vectors;

b) $\mathbb{E}\, X_{12} = \mathbb{E}\, X_{21} = 0, \mathbb{E}\, X_{12}^2 = \mathbb{E}\, X_{21}^2 = 1$ and $\max(\mathbb{E}\, |X_{12}|^4, \mathbb{E}\, |X_{21}|^4) \leq M_4$;

c) $\mathbb{E}(X_{12}X_{21}) = \rho, |\rho| \leq 1$;

d) The diagonal entries X_{ii} are i.i.d. random variables, independent of off-diagonal entries, $\mathbb{E}\, X_{11} = 0$ and $\mathbb{E}\, X_{11}^2 < \infty$.

Denote by $\lambda_1, ..., \lambda_n$ the eigenvalues of the matrix $n^{-1/2}\mathbf{X}_n$ and define empirical spectral measure by

$$\mu_n(B) = \frac{1}{n}\#\{1 \leq i \leq n : \lambda_i \in B\}, \quad B \in \mathcal{B}(\mathbb{C}),$$

where $\mathcal{B}(\mathbb{C})$ is a Borel σ-algebra of $\mathbb{C}$.

We say that the sequence of random probability measures $m_n(\cdot)$ converges weakly in probability to probability measure $m(\cdot)$ if for all continues and bounded functions $f : \mathbb{C} \to \mathbb{C}$ and all $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int_{\mathbb{C}} f(x)m_n(dz) - \int_{\mathbb{C}} f(x)m(dz)\right| > \varepsilon\right) = 0.$$

We denote weak convergence by symbol $\xrightarrow{weak}$.

A fundamental problem in the theory of random matrices is to determine the limiting distribution of μ_n as the size of the random matrix tends to infinity. The main result of this paper is the following

Theorem 1.1. (Elliptic Law) *Let $\mathbf{X}_n$ satisfies condition (**C0**) and $|\rho| < 1$. Then $\mu_n \xrightarrow{weak} \mu$ in probability, and μ has a density g:*

$$g(x,y) = \begin{cases} \frac{1}{\pi(1-\rho^2)}, & x,y \in \mathcal{E}, \\ 0, & otherwise, \end{cases}$$

where

$$\mathcal{E} := \left\{x,y \in \mathbb{R} : \frac{x^2}{(1+\rho)^2} + \frac{y^2}{(1-\rho)^2} \leq 1\right\}.$$

Theorem 1.1 asserts that under assumption of finite fourth moment empirical distribution weakly converges in probability to uniform distribution on the ellipse. The axis of the ellipse are determined by correlation between X_{12} and X_{21}. This result was called by Girko "Elliptic Law". Limit distribution doesn't depend on distribution of matrix elements and the result in this sense is universal.

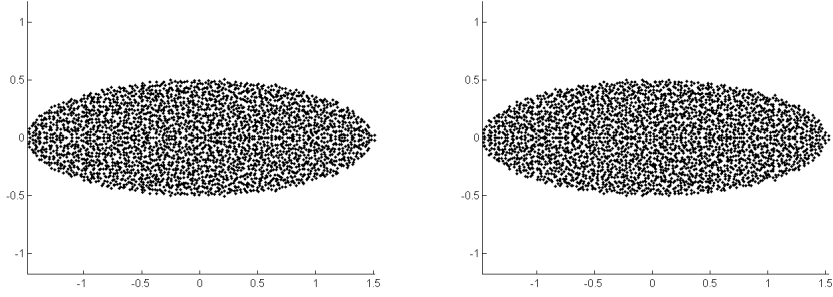Figure 1 illustates Elliptic law for $\rho = 0.5$ and Figure 2 – for $\rho = -0.5$.

FIGURE 1. Eigenvalues of matrix $n^{-1/2}\mathbf{X}$ for $n = 3000$ and $\rho = 0.5$. On the left, each entry is an iid Gaussian normal random variable. On the right, each entry is an iid Bernoulli random variable, taking the values $+1$ and -1 each with probability $1/2$.
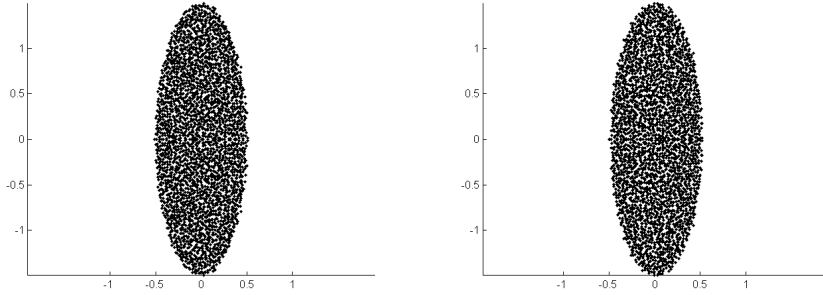

FIGURE 2. Eigenvalues of matrix $n^{-1/2}\mathbf{X}$ for $n = 3000$ and $\rho = -0.5$. On the left, each entry is an iid Gaussian normal random variable. On the right, each entry is an iid Bernoulli random variable, taking the values $+1$ and -1 each with probability $1/2$.

In 1985 Girko proved elliptic law for rather general ensembles of random matrices under assumption that matrix elements have a density, see [7] and [8]. Girko used method of characteristic functions. Using V-transform he reduced problem to the problem for Hermitian matrices $(n^{-1/2}\mathbf{X}_n - z\mathbf{I})^*(n^{-1/2}\mathbf{X}_n - z\mathbf{I})$ and established convergence of empirical spectral distribution of singular values of $n^{-1/2}\mathbf{X}_n - z\mathbf{I}$ to the limit which determines the elliptic law.

Let elements of real asymmetric random matrix $\mathbf{X}$ have Gaussian distribution with zero mean and correlations

$$\mathbb{E}\, X_{ij}^2 = 1 \text{ and } \mathbb{E}\, X_{ij}X_{ij} = \rho, \quad i \neq j, \quad |\rho| < 1.$$

The ensemble of such matrices can be specified by the probability measure

$$\mathbb{P}(dX) \sim \exp\left[-\frac{n}{2(1-\rho^2)}\operatorname{Tr}(XX^T - \rho X^2)\right].$$

It was proved that $\mu_n \xrightarrow{weak} \mu$, where μ has a density from Theorem 1.1, see [14]. We will use this result to prove Theorem 1.1 in the general case.

Remark 1.2. *This result can be generalized to an ensemble of Gaussian complex asymmetric matrices. In this case, the invariant measure is*

$$\mathbb{P}(dX) \sim \exp\left[-\frac{n}{1-|\rho|^2}\operatorname{Tr}(XX^T - 2\operatorname{Re}\rho X^2)\right]$$

and $\mathbb{E}\,|X_{ij}|^2 = 1, \mathbb{E}\,X_{ij}X_{ji} = |\rho|e^{2i\theta}$ for $i \neq j$. Then the limit measure has a uniform density inside an ellipse which is centered at zero and has semiaxes $1 + |\rho|$ in the direction θ and $1 - |\rho|$ in the direction $\theta + \pi/2$.

For the discussion of elliptic law in Gaussian case see also [6], [1, Chapter 18] and [10].

We repeat physical motivation of models of random matrices which satisfy condition (**C0**) from [14]: *"The statistical properties of random asymmetric matrices may be important in the understanding of the behavior of certain dynamical systems far from equilibrium. One example is the dynamics of neural networks. A simple dynamic model of neural network consists of n continues "scalar" degrees of freedom("neurons") obeying coupled nonlinear differential equations ("circuit equations"). The coupling between the neurons is given by a synaptic matrix $\mathbf{X}$ which, in general, is asymmetric and has a substantial degree of disorder. In this case, the eigenstates of the synaptic matrix play an important role in the dynamics particulary when the neuron nonlinearity is not big".*

It will be interesting to prove Theorem 1.1 only under assumption of finite second moment and prove sparse analogs. It is the direction of our further research.

If $\rho = 0$ we assume that all entries of $\mathbf{X}_n$ are independent random variables and Circular law holds (see [2], [16], [9]):

Theorem 1.3. (Circular law) *Let $\mathbf{X}_n$ be a random matrix with independent identically distributed entries, $\mathbb{E}\,X_{ij} = 0$ and $\mathbb{E}\,X_{ij}^2 = 1$. Then $\mu_n \xrightarrow{weak} \mu$ in probability, and μ has uniform density on the unit circular.*

See Figure 3 for illustration of Circulaw law.

If $\rho = 1$ then matrix $\mathbf{X}_n$ is symmetric and its eigenvalues are real numbers. In this case the next theorem is known as a Wigner's semi-circular law (see [2]):

Theorem 1.4. (Semi-circular law) *Let $\mathbf{X}_n$ be a symmetric random matrix with independent identically distributed entries for $i \geq j$, $\mathbb{E}\,X_{ij} = 0$, $\mathbb{E}\,X_{ij}^2 = 1$. Then $\mu_n \xrightarrow{weak} \mu$ in probability, and μ has a density g:*

$$g(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2}, & -2 \leq x \leq 2, \\ 0, & otherwise. \end{cases}$$
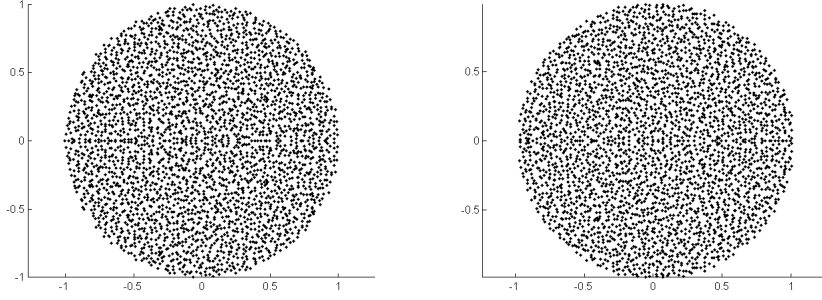
FIGURE 3. Eigenvalues of matrix $n^{-1/2}\mathbf{X}$ for $n = 3000$ and $\rho = 0$. On the left, each entry is an iid Gaussian normal random variable. On the right, each entry is an iid Bernoulli random variable, taking the values $+1$ and -1 each with probability $1/2$.

Throughout this paper we assume that all random variables are defined on common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we will write almost surely (a.s) instead of $\mathbb{P}$-almost surely. By $\mathrm{Tr}(\mathbf{A})$ and $\mathrm{rk}(\mathbf{A})$ we mean trace and rank of the matrix $\mathbf{A}$ respectively. We denote singular values of matrix $\mathbf{A}$ by $s_i(\mathbf{A})$ and $s_1(\mathbf{A}) \geq s_2(\mathbf{A}) \geq ... \geq s_n(\mathbf{A})$. For vector $x = (x_1, ..., x_n)$ we introduce $||x||_2 := (\sum_{i=1}^{n} x_i^2)^{1/2}$ and $||x||_3 := (\sum_{i=1}^{3} |x_i|^3)^{1/3}$. We denote unit sphere and unit ball by $S^{n-1} := \{x : ||x||_2 = 1\}$ and $B_1^n := \{x : ||x||_2 \leq 1\}$ respectively. For matrix $\mathbf{A}$ define spectral norm by $||\mathbf{A}|| := \sup_{x:||x||_2=1} ||\mathbf{A}x||_2$ and Hilbert-Schmidt norm by $||\mathbf{A}||_{HS} := (\mathrm{Tr}(\mathbf{A}^*\mathbf{A}))^{1/2}$. By $[n]$ we mean the set $\{1, ..., n\}$. We denote by $\mathcal{B}(\mathbb{T})$ - Borel σ-algebra of $\mathbb{T}$, where $\mathbb{T} = \mathbb{R}$ or $\mathbb{C}$.

2. PROOF OF THE MAIN RESULT

Further we will need the definition of logarithmic potential (see [12]) and uniform integrability of function with respect to the sequence of probability measures.

Definition 2.1. *The logarithmic potential U_m of measure $m(\cdot)$ is a function $U_m : \mathbb{C} \to (-\infty, +\infty]$ defined for all $z \in \mathbb{C}$ by*

$$U_m(z) = -\int_{\mathbb{C}} \log |z - w| m(dw).$$

Definition 2.2. *The function $f : \mathbb{T} \to \mathbb{R}$, where $\mathbb{T} = \mathbb{C}$ or $\mathbb{T} = \mathbb{R}$, is uniformly integrable in probability with respect to the sequence of random measures $\{m_n\}_{n\geq 1}$ on $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ if for all $\varepsilon > 0$:*

$$\lim_{t \to \infty} \overline{\lim_{n \to \infty}} \mathbb{P}\left(\int_{|f|>t} |f(x)| m_n(dx) > \varepsilon\right) = 0.$$

Let $s_1(n^{-1/2}\mathbf{X} - z\mathbf{I}) \geq s_2(n^{-1/2}\mathbf{X} - z\mathbf{I}) \geq ... \geq s_n(n^{-1/2}\mathbf{X} - z\mathbf{I})$ be singular values of $n^{-1/2}\mathbf{X}_n - z\mathbf{I}$ and

$$\nu_n(z, B) = \frac{1}{n}\#\{i \geq 1 : s_i(n^{-1/2}\mathbf{X} - z\mathbf{I}) \in B\}, \quad B \in \mathcal{B}(\mathbb{R})-$$

empirical spectral measure of singular values. We will omit argument z in notation of measure $\nu_n(z, B)$ if it doesn't confuse.

The convergence in the Theorem 1.1 will be proved via convergence of logarithmic potential of μ_n to the logarithmic potential of μ. We can rewrite logarithmic potential of μ_n via the logarithmic moments of measure ν_n by

$$U_{\mu_n}(z) = -\int_{\mathbb{C}} \log|z - w|\mu_n(dw) = -\frac{1}{n}\log\left|\det\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - z\mathbf{I}\right)\right|$$

$$= -\frac{1}{2n}\log\det\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - z\mathbf{I}\right)^*\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - z\mathbf{I}\right) = -\int_0^\infty \log x\nu_n(dx).$$

This allows us to consider Hermitian matrix $(n^{-1/2}\mathbf{X}_n - z\mathbf{I})^*(n^{-1/2}\mathbf{X}_n - z\mathbf{I})$ instead of asymmetric $n^{-1/2}\mathbf{X}$. To prove Theorem 1.1 we need the following

Lemma 2.3. *Let $(\mathbf{X}_n)_{n\geq1}$ be a sequence of $n \times n$ random matrices. Suppose that for a.a. $z \in \mathbb{C}$ there exists a probability measure ν_z on $[0, \infty)$ such that*
a) $\nu_n \xrightarrow{weak} \nu_z$ as $n \to \infty$ in probability
b) $\log$ is uniformly integrable in probability with respect to $\{\nu_n\}_{n\geq1}$.

Then there exists a probability measure μ such that
a) $\mu_n \xrightarrow{weak} \mu$ as $n \to \infty$ in probability
b) for a.a. $z \in \mathbb{C}$
$$U_\mu(z) = -\int_0^\infty \log x\nu_z(dx).$$

Proof. See [3, Lemma 4.3] for the proof. $\square$

Proof. (Proof of Theorem 1.1) Our aim is to prove convergence of ν_n to ν_z, uniform integrability of $\log(\cdot)$ with respect to $\{\nu_n\}_{n\geq1}$ and show that ν_z determines elliptic law.

From Theorem 4.1 we can conclude uniform integrability of $\log(\cdot)$. The proof of Theorem 4.1 is based on Theorem 3.1 and some additional results.

In Theorem 5.2 it is proved that $\nu_n \xrightarrow{weak} \nu_z$ in probability, where ν_z is some probability measure, which doesn't depend on distribution of elements of matrix $\mathbf{X}$.

If matrix $\mathbf{X}$ has Gaussian elements we redenote μ_n by $\hat{\mu}_n$.

By Lemma 2.3 there exists probability measure $\hat{\mu}$ such that $\mu_n \xrightarrow{weak} \hat{\mu}$ in probability and $U_{\hat{\mu}}(z) = -\int_0^\infty \log x\nu_z(dx)$. But in Gaussian case $\mu_n \xrightarrow{weak} \mu$ in

probability and $U_\mu(z) = -\int_0^\infty \log x \nu_z(dx)$. We know that ν_z is the same for all matrices which satisfy condition (**C0**) and we have

$$U_{\hat\mu}(z) = -\int_0^\infty \log x \nu_z(dx) = U_\mu(z).$$

From unicity of logarithmic potential we conclude that $\hat\mu = \mu$. $\qquad\square$

3. Least singular value

From properties of the largest and the smallest singular values

$$s_1(\mathbf{A}) = ||\mathbf{A}|| = \sup_{x:||x||_2=1} ||\mathbf{A}x||_2, \quad s_n(\mathbf{A}) = \inf_{x:||x||_2=1} ||\mathbf{A}x||_2.$$

To prove uniform integrability of $\log(\cdot)$ we need to estimate probability of the event $\{s_n(\mathbf{A}) \le \varepsilon n^{-1/2}, ||\mathbf{X}|| \le K\sqrt{n}\}$, where $\mathbf{A} = \mathbf{X} - z\mathbf{I}$. We can assume that $\varepsilon n^{-1/2} \le K n^{1/2}$. If $|z| \ge 2K\sqrt{n}$ then probability of the event is automatically zero. So we can consider the case when $|z| \le 2K n^{1/2}$. We have $||\mathbf{A}|| \le ||\mathbf{X}|| + |z| \le 3K n^{1/2}$. In this section we prove theorem

Theorem 3.1. *Let* $\mathbf{A} = \mathbf{X} - z\mathbf{I}$, *where* $\mathbf{X}$ *is* $n \times n$ *random matrix satisfying* (**C0**). *Let* $K > 1$. *Then for every* $\varepsilon > 0$ *one has*

$$\mathbb{P}(s_n(\mathbf{A}) \le \varepsilon n^{-1/2}, ||\mathbf{A}|| \le 3K\sqrt{n}) \le C(\rho)\varepsilon^{1/8} + C_1(\rho)n^{-1/8},$$

where $C(\rho), C_1(\rho)$ *are some constants which can depend only on* ρ, K *and* M_4.

Remark 3.2. *Mark Rudelson and Roman Vershynin in [11] and Roman Vershynin in [17] found bounds for the least singular value of matrices with independent entries and symmetric matrices respectively. In this section we will follow their ideas.*

3.1. The small ball probability via central limit theorem. We recall definition of Levy concentration function

Definition 3.3. *Levy concentration function of random variable* Z *with values from* $\mathbb{R}^d$ *is a function*

$$\mathcal{L}(Z, \varepsilon) = \sup_{v \in \mathbb{R}^d} \mathbb{P}(||Z - v||_2 < \varepsilon).$$

The next statement gives the bound for Levy concentration function of sum of independent random variables in $\mathbb{R}$.

Statement 3.4. *Let* $\{a_i\xi_i + b_i\eta_i\}_{i\ge1}$ *be independent random variables,* $\mathbb{E}\,\xi_i = \mathbb{E}\,\eta_i = 0$, $\mathbb{E}\,\xi_i^2 \ge 1, \mathbb{E}\,\eta_i^2 \ge 1$, $\mathbb{E}\,\xi_i\eta_i = \rho$, $\max(\mathbb{E}\,\xi_i^4, \mathbb{E}\,\eta_i^4) \le M_4$, $a_i^{-1}b_i = O(1)$. *We assume that* $\tau(2n)^{-1/2} \le |a_i| \le (\delta n)^{-1/2}$, *where* δ, τ *are some constants. Then*

$$\mathcal{L}\left(\sum_{i=1}^n (a_i\xi_i + b_i\eta_i), \varepsilon\right) \le \frac{C\varepsilon}{(1-\rho^2)^{1/2}} + \frac{C_1}{(1-\rho^2)^{3/2}n^{1/2}}.$$

Proof. Set $\sigma_{i1}^2 = \mathbb{E}\,\xi_i^2$ and $\sigma_{i2}^2 = \mathbb{E}\,\eta_i^2$. It is easy to see that

$$\sigma^2 = \mathbb{E}(\sum_{i=1}^n Z_i)^2 =$$

$$= \sum_{i=1}^n |a_i|^2 (\sigma_{i1}^2 + 2\rho\sigma_{i1}\sigma_{i2}a_i^{-1}b_i + \sigma_{i2}^2(a_i^{-1}b_i)^2) \geq (1-\rho^2)\sum_{i=1}^n \sigma_{i1}^2 |a_i|^2$$

and

$$\sum_{i=1}^n \mathbb{E}\,|a_i\xi_i + b_i\eta_i|^3 \leq \sum_{i=1}^n |a_i|^3 \mathbb{E}\,|\xi_i + a_i^{-1}b_i\eta_i|^3 \leq C'M_4\|a\|_3^3,$$

where we have used the fact $a_i^{-1}b_i = O(1)$. By Central Limit Theorem A.1 for arbitrary vector $v \in \mathbb{R}$

$$\mathbb{P}\left(\left|\sum_{i=1}^n (a_i\xi_i + b_i\eta_i) - v\right| \leq \varepsilon\right) \leq \mathbb{P}\left(|g' - v| \leq \varepsilon\right) + C''\frac{\sum_{i=1}^n \mathbb{E}\,|a_i\xi_i + b_i\eta_i|^3}{\sigma^3},$$

where g' has gaussian distribution with zero mean and variance σ^2. The density of g' is uniformly bounded by $1/\sqrt{2\pi\sigma^2}$. We have

$$\mathbb{P}\left(|\sum_{i=1}^n (a_i\xi_i + b_i\eta_i) - v| \leq \varepsilon\right) \leq \frac{C\varepsilon}{(1-\rho^2)^{1/2}} + \frac{C_1}{(1-\rho^2)^{3/2}n^{1/2}}.$$

We can take maximum and conclude the statement. $\qquad\square$

Remark 3.5. *Let us consider the case $b_i = 0$ for all $i \geq 1$. It is easy to show that*

$$\mathcal{L}\left(\sum_{i=1}^n a_i\xi_i, \varepsilon\right) \leq C(\varepsilon + n^{-1/2}).$$

3.2. **Decomposition of the sphere and invertibility.** To prove Theorem 3.1, we shall partition the unit sphere S^{n-1} into the two sets of compressible and incompressible vectors, and show the invertibility of **A** on each set separately.

Definition 3.6. *(Compressible and incompressible vectors) Let $\delta, \tau \in (0,1)$. A vector $x \in \mathbb{R}^n$ is called sparse if $|supp(x)| \leq \delta n$. A vector $x \in S^{n-1}$ is called compressible if x is within Euclidian distance τ from the set of all sparse vectors. A vector $x \in S^{n-1}$ is called incompressible if it is not compressible. The sets of sparse, compressible and incompressible vectors will be denoted by Sparse $=$ Sparse (δ), Comp $=$ Comp (δ, τ) and Incomp $=$ Incomp(δ, τ) respectively.*

We first estimate $\|\mathbf{A}x\|$ for a fixed vector $x \in S^{n-1}$. The next statement can be found in [17]

Lemma 3.7. *Let $\mathbf{A}$ be a matrix from Theorem 3.1 and let $K > 1$. There exist constants $\delta, \tau, c \in (0,1)$ that depend only on K and M_4 and such that the following holds. For every $u \in \mathbb{R}^n$, one has*

$$(3.1) \quad \mathbb{P}\left(\inf_{\frac{x}{\|x\|_2} \in Comp(\delta,\tau)} \|\mathbf{A}x - u\|_2/\|x\|_2 \leq c_4\sqrt{n}, \|\mathbf{A}\| \leq 3K\sqrt{n}\right) \leq 2e^{-cn}.$$

Proof. See [17, Statement 4.2]. The proof of this result for matrices which satisfy condition (**C0**) can be carried out by similar arguments. $\square$

For the incompressible vectors, we shall reduce the invertibility problem to a lower bound on the distance between a random vector and a random hyperplane. For this aim we recall Lemma 3.5 from [11]

Lemma 3.8. *Let* **A** *be a random matrix from theorem. Let* $A_1, ..., A_n$ *denote the column vectors of* **A***, and let* H_k *denote the span of all columns except the* k*-th. Then for every* $\delta, \tau \in (0, 1)$ *and every* $\varepsilon > 0$*, one has*

$$(3.2) \qquad \mathbb{P}\left(\inf_{x \in Incomp(\delta, \tau)} ||\mathbf{A}x||_2 < \varepsilon n^{-1}\right) \leq \frac{1}{\delta n} \sum_{k=1}^{n} \mathbb{P}(\text{dist}(A_k, H_k) < \tau^{-1}\varepsilon).$$

Lemma 3.8 reduces the invertibility problem to a lower bound on the distance between a random vector and a random hyperplane.

We decompose matrix $\mathbf{A} = \mathbf{X} - z\mathbf{I}$ into the blocks

$$(3.3) \qquad \begin{pmatrix} a_{11} & V^T \\ U & \mathbf{B} \end{pmatrix}$$

where **B** is $(n-1) \times (n-1)$ matrix, $U, V \in \mathbb{R}^{n-1}$.

Let h be any unit vector orthogonal to $A_2, ..., A_n$. It follows that

$$0 = \begin{pmatrix} V^T \\ \mathbf{B} \end{pmatrix}^T h = h_1 V + \mathbf{B}^T g,$$

where $h = (h_1, g)$, and

$$g = -h_1 \mathbf{B}^{-T} V$$

From definition of h

$$1 = ||h||_2^2 = |h_1|^2 + ||g||_2^2 = |h_1|^2 + |h_1|^2 ||\mathbf{B}^{-T}V||_2^2$$

Using this equations we estimate distance

$$\text{dist}(A_1, H) \geq |(A_1, h)| = \frac{|a_{11} - (\mathbf{B}^{-T}V, U)|}{\sqrt{1 + ||\mathbf{B}^{-T}V||_2^2}}$$

It is easy to show that $||\mathbf{B}|| \leq ||\mathbf{A}||$. Let vector $e_1 \in S^{n-2}$ be such that $||\mathbf{B}|| = ||\mathbf{B}e_1||_2$. Then we can take vector $e = (0, e_1)^T \in S^{n-1}$ and for this vector

$$||\mathbf{A}|| \geq ||\mathbf{A}e||_2 = ||(V^T e_1, \mathbf{B}e_1)^T||_2 \geq ||\mathbf{B}e_1||_2 = ||\mathbf{B}||.$$

The bound for right hand sand of (3.2) will follow from the

Lemma 3.9. *Let matrix* **A** *be from Theorem 3.1. Then for all* $\varepsilon > 0$

$$(3.4) \quad \sup_{v \in \mathbb{R}} \mathbb{P}\left(\frac{|(\mathbf{B}^{-T}V, U) - v|}{\sqrt{1 + ||\mathbf{B}^{-T}V||_2^2}} \leq \varepsilon, ||\mathbf{B}|| \leq 3K\sqrt{n}\right) \leq C(\rho)\varepsilon^{1/8} + C'(\rho)n^{-1/8},$$

where $\mathbf{B}, U, V$ *are determined by (3.3) and* $C(\rho), C_1(\rho)$ *are some constants which can depend only on* ρ, K *and* M_4.

To get this bound we need several statements. We introduce matrix

$$(3.5) \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{O}_{n-1} & \mathbf{B}^{-T} \\ \mathbf{B}^{-1} & \mathbf{O}_{n-1} \end{pmatrix} \quad W = \begin{pmatrix} U \\ V \end{pmatrix},$$

where $\mathbf{O}_{n-1}$ is $(n-1) \times (n-1)$ matrix with zero entries. Scalar product in (3.4) can be rewritten using definition of Q:

$$(3.6) \qquad \sup_{v \in \mathbb{R}} \mathbb{P} \left(\frac{|(\mathbf{Q}W, W) - v|}{\sqrt{1 + ||\mathbf{B}^{-T}V||_2^2}} \le 2\varepsilon \right).$$

Introduce vectors

$$(3.7) \qquad W' = \begin{pmatrix} U' \\ V' \end{pmatrix} \quad Z = \begin{pmatrix} U \\ V' \end{pmatrix},$$

where U', U' are independent copies of U, V respectively. We need the following

Statement 3.10.

$$\sup_{v \in \mathbb{R}} \mathbb{P}_W \left(|(\mathbf{Q}W, W) - v| \le 2\varepsilon \right) \le \mathbb{P}_{W,W'} \left(|(\mathbf{Q}\mathbf{P}_{J^c}(W - W'), \mathbf{P}_J W) - u| \le 2\varepsilon \right),$$

where u doesn't depend on $\mathbf{P}_J W = (\mathbf{P}_J U, \mathbf{P}_J V)^T$.

Proof. Let us fix v and denote

$$p := \mathbb{P} \left(|(\mathbf{Q}W, W) - v| \le 2\varepsilon \right).$$

We can decompose the set $[n]$ into union $[n] = J \cup J^c$. We can take $U_1 = \mathbf{P}_J U, U_2 = \mathbf{P}_{J^c} U, V_1 = \mathbf{P}_J V$ and $V_2 = \mathbf{P}_{J^c} V$. By Lemma A.2

$$(3.8) \qquad p^2 \le \mathbb{P} \left(|(\mathbf{Q}W, W) - v| \le 2\varepsilon, |(\mathbf{Q}Z, Z) - v| \le 2\varepsilon \right)$$
$$\le \mathbb{P} \left(|(\mathbf{Q}W, W) - (\mathbf{Q}Z, Z)| \le 4\varepsilon \right).$$

Let us rewrite $\mathbf{B}^{-T}$ in the block form

$$\mathbf{B}^{-T} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}.$$

We have

$$(\mathbf{Q}W, W) = (\mathbf{E}V_1, U_1) + (\mathbf{F}V_2, U_1) + (\mathbf{G}V_1, U_2) + (\mathbf{H}V_2, U_2)$$
$$+ (\mathbf{E}^T U_1, V_1) + (\mathbf{G}^T U_2, V_1) + (\mathbf{F}^T U_1, V_2) + (\mathbf{H}^T U_2, V_2)$$
$$(\mathbf{Q}Z, Z) = (\mathbf{E}V_1, U_1) + (\mathbf{F}V_2', U_1) + (\mathbf{G}V_1, U_2') + (\mathbf{H}V_2', U_2')$$
$$+ (\mathbf{E}^T U_1, V_1) + (\mathbf{G}^T U_2', V_1) + (\mathbf{F}^T U_1, V_2') + (\mathbf{H}^T U_2', V_2')$$

and

$$(3.9) \qquad (\mathbf{Q}W, W) - (\mathbf{Q}Z, Z) = 2(\mathbf{F}(V_2 - V_2'), U_1) + 2(\mathbf{G}^T(U_2 - U_2'), V_1)$$
$$+ 2(\mathbf{H}V_2, V_2) - 2(\mathbf{H}V_2', V_2').$$

The last two terms in (3.9) depend only on U_2, U_2', V_2, V_2' and we conclude that

$$p_1^2 \le \mathbb{P} \left(|(\mathbf{Q}P_{J^c}(W - W'), P_J W) - u| \le 2\varepsilon \right),$$

where $u = u(U_2, V_2, U_2', V_2', \mathbf{F}, \mathbf{G}, \mathbf{H})$. $\qquad \square$

Statement 3.11. *For all $u \in \mathbb{R}^{n-1}$*

$$\mathbb{P}\left(\frac{\mathbf{B}^{-T}u}{||\mathbf{B}^{-T}u||_2} \in Comp(\delta, \tau) \ \text{and} \ ||\mathbf{B}|| \leq 3Kn^{1/2}\right) \leq 2e^{-cn}.$$

Proof. Let $x = \mathbf{B}^{-T}u$. It is easy to see that

$$\left\{\frac{\mathbf{B}^{-T}u}{||\mathbf{B}^{-T}u||_2} \in Comp(\delta, \tau)\right\} \subseteqq \left\{\exists x : \frac{x}{||x||_2} \in Comp(\delta, \tau) \ \text{and} \ \mathbf{B}^T x = u\right\}$$

Replacing matrix $\mathbf{A}$ with $\mathbf{B}^T$ one can easily check that the proof of Lemma 3.7 remains valid for $\mathbf{B}^T$ as well as for $\mathbf{A}$. $\qquad\qquad\square$

Remark 3.12. *The Statement 3.11 holds true for $\mathbf{B}^{-T}$ replaced with $\mathbf{B}^{-1}$.*

Statement 3.13. *Let $\mathbf{A}$ satisfies condition (**C0**) and $\mathbf{B}$ be a matrix from decomposition (3.3). Assume that $||\mathbf{B}|| \leq 3K\sqrt{n}$. Then with probability at least $1 - e^{-cn}$ matrix $\mathbf{B}$ has the following properties:*

    a) *$||\mathbf{B}^{-T}V||_2 \geq C$ with probability $1 - e^{-c'n}$ in W,*

    b) *$||\mathbf{B}^{-T}V||_2 \leq \varepsilon^{-1/2}||\mathbf{B}^{-T}||_{HS}$ with probability $1 - \varepsilon$ in V,*

    c) *$||\mathbf{Q}W||_2 \geq \varepsilon||\mathbf{B}^{-T}||_{HS}$ with probability $1 - C'(\varepsilon + n^{-1/2})$ in W.*

Proof. Let $\{e_k\}_{k=1}^n$ be a standard basis in $\mathbb{R}^n$. For all $1 \leq k \leq n$ define vectors by

$$x_k := \frac{\mathbf{B}^{-1}e_k}{||\mathbf{B}^{-1}e_k||_2}.$$

By Statement 3.11 vector x_k is incompressible with probability $1 - e^{-cn}$. We fix matrix $\mathbf{B}$ with such property.
a) By norm inequality $||V||_2 \leq ||\mathbf{B}||||\mathbf{B}^{-T}V||_2$. We know that $||\mathbf{B}|| \leq 3K\sqrt{n}$. By Lemma A.7 and Lemma A.9 $||V||_2 \geq \sqrt{n}$. So we have that $||\mathbf{B}^{-1}V||_2 \geq C$ with probability $1 - e^{-c'n}$.
b) By definition

$$||\mathbf{B}^{-T}V||_2^2 = \sum_{i=1}^n (\mathbf{B}^{-1}e_k, V)^2 = \sum_{i=1}^n ||\mathbf{B}^{-1}e_i||_2^2 (x_k, V)^2.$$

It is easy to see that $\mathbb{E}(V, x_k)^2 = 1$. So

$$\mathbb{E}||\mathbf{B}^{-T}V||_2^2 = \sum_{i=1}^n ||\mathbf{B}^{-1}e_i||_2^2 = ||\mathbf{B}^{-1}||_{HS}^2.$$

By Markov inequality

$$\mathbb{P}(||\mathbf{B}^{-T}V||_2 \geq \varepsilon^{-1/2}||\mathbf{B}^{-1}||_{HS}) \leq \varepsilon.$$

c) By Lemma A.3, Lemma A.4, Lemma A.6 and Remark 3.5

$$\mathbb{P}(||\mathbf{Q}W||_2 \leq \varepsilon ||\mathbf{B}^{-1}||_{HS}) \leq \mathbb{P}(||\mathbf{B}^{-T}V||_2 \leq \varepsilon ||B^{-1}||_{HS})$$

$$= \mathbb{P}(||\mathbf{B}^{-T}V||_2^2 \leq \varepsilon ||\mathbf{B}^{-1}||_{HS}^2) = \mathbb{P}(\sum_{i=1}^{n} ||\mathbf{B}^{-1}e_i||_2 (x_i, V)^2 \leq \varepsilon^2 ||\mathbf{B}^{-1}||_{HS}^2)$$

$$= \mathbb{P}(\sum_{i=1}^{n} p_i (x_i, V)^2 \leq \varepsilon^2) \leq 2 \sum_{i=1}^{n} p_i \, \mathbb{P}((x_i, V) \leq \sqrt{2}\varepsilon) \leq C'(\varepsilon + n^{-1/2}).$$

$\square$

Proof. (proof of Lemma 3.9) Let $\xi_1, ..., \xi_n$ be i.i.d. Bernoulli random variables with $\mathbb{E}\,\xi_i = c_0/2$. We define $J := \{i : \xi_i = 0\}$ and $\mathbb{E}_0 := \{|J^c| \leq c_0 n\}$. From large deviation inequality we may conclude that $\mathbb{P}(E_0) \geq 1 - 2\exp(-c_0^2 n/2)$. Introduce event

$$E_1 := \{\varepsilon_0^{1/2} \sqrt{1 + ||\mathbf{B}^{-T}V||_2^2} \leq ||\mathbf{B}^{-1}||_{HS} \leq \varepsilon_0^{-1} ||\mathbf{Q}\mathbf{P}_{J^c}(W - W')||_2\},$$

where ε_0 will be choosen later.

From Statement 3.13 we can conclude that

$$\mathbb{P}_{\mathbf{B}, W, W', J}(E_1 \cup ||B|| \geq 3K\sqrt{n}) \geq 1 - C'(\varepsilon_0 + n^{-1/2}) - 2e^{-c'n}.$$

Consider the random vector

$$w_0 = \frac{1}{||\mathbf{Q}\mathbf{P}_{J^c}(W - W')||_2} \begin{pmatrix} \mathbf{B}^{-T}\mathbf{P}_{J^c}(V - V') \\ \mathbf{B}^{-1}\mathbf{P}_{J^c}(U - U') \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

By Statement 3.11 it follows that the event $E_2 := \{a \in incomp(\delta, \tau)\}$ holds with probability

$$\mathbb{P}_{\mathbf{B}}(E_2 \cup ||\mathbf{B}|| \geq 3K\sqrt{n}|W, W', J) \geq 1 - 2\exp(-c''n).$$

Combining these probabilities we have

$$\mathbb{P}_{\mathbf{B}, W, W', J}(E_0, \ E_1, \ E_2 \cup ||\mathbf{B}|| \geq 3K\sqrt{n})$$

$$\geq 1 - 2e^{-c_0^2 n/2} - C'(\varepsilon_0 + n^{-1/2}) - 2e^{-c'n} - 2e^{-c''n} := 1 - p_0.$$

We may fix J that satisfies $|J^c| \leq c_0$ and

$$\mathbb{P}_{\mathbf{B}, W, W'}(E_1, \ E_2 \cup ||\mathbf{B}|| \geq 3K\sqrt{n}) \geq 1 - p_0.$$

By Fubini's theorem $\mathbf{B}$ has the following property with probability at least $1 - \sqrt{p_0}$

$$\mathbb{P}_{W, W'}(E_1, \ E_2 \cup ||\mathbf{B}|| \geq 3K\sqrt{n}|\mathbf{B}) \geq 1 - \sqrt{p_0}.$$

The event $\{||\mathbf{B}|| \geq 3K\sqrt{n}\}$ depends only on $\mathbf{B}$. We may conclude that random matrix $\mathbf{B}$ has the following property with probability at least $1 - \sqrt{p_0}$: either $||\mathbf{B}|| \geq 3K\sqrt{n}$, or

$$(3.10) \qquad ||\mathbf{B}|| \leq 3K\sqrt{n} \text{ and } \mathbb{P}_{W, W'}(E_1, \ E_2|\mathbf{B}) \geq 1 - \sqrt{p_0}$$

The event we are interested in is

$$\Omega_0 := \left(\frac{|(\mathbf{Q}W, W) - u|}{\sqrt{1 + ||\mathbf{B}^{-T}V||_2^2}} \leq 2\varepsilon \right).$$

We need to estimate probability

$$\mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap ||\mathbf{B}|| \leq 3K\sqrt{n}) \leq \mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap (3.10) \text{ holds}) + \mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap (3.10) \text{ fails}).$$

The last term is bounded by $\sqrt{p_0}$.

$$\mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap ||\mathbf{B}|| \leq 3K\sqrt{n}) \leq \sup_{\mathbf{B} \text{ satisfies } (3.10)} \mathbb{P}_W(\Omega_0|\mathbf{B}) + \sqrt{p_0}.$$

We can conclude that

$$\mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap ||\mathbf{B}|| \leq 3K\sqrt{n}) \leq \sup_{\mathbf{B} \text{ satisfies } (3.10)} \mathbb{P}_{W,W'}(\Omega_0, \ E_1|\mathbf{B}) + 2\sqrt{p_0}.$$

Let us fix $\mathbf{B}$ that satisfies (3.10) and denote $p_1 := \mathbb{P}_{W,W'}(\Omega_0, \ E_1|\mathbf{B})$. By Statement 3.10 and the first inequality in E_1 we have

$$p_1^2 \leq \mathbb{P}_{W,W'}\left(\underbrace{|(\mathbf{Q}\mathbf{P}_{J^c}(W - W'), \mathbf{P}_J W) - v| \leq \frac{\varepsilon}{\sqrt{\varepsilon_0}}||\mathbf{B}^{-1}||_{HS}}_{\Omega_1}\right)$$

and

$$\mathbb{P}_{W,W'}(\Omega_1) \leq \mathbb{P}_{W,W'}(\Omega_1, \ E_1, \ E_2) + \sqrt{p_0}.$$

Further

$$p_1^2 \leq \mathbb{P}_{W,W'}(|(w_0, \mathbf{P}_J W) - v| \leq 2\varepsilon_0^{-3/2}\varepsilon, \ E_2) + \sqrt{p_0}.$$

By definition random vector w_0 is determined by the random vector $\mathbf{P}_{J^c}(W - W')$, which is independent of the random vector $\mathbf{P}_J W$. We fix $\mathbf{P}_{J^c}(W - W')$ and have

$$p_1^2 \leq \sup_{\substack{w_0=(a,b)^T: \\ a \in Incomp(\delta,\tau) \\ w \in \mathbb{R}}} \mathbb{P}_{\mathbf{P}_J W}\left(|(w_0, \mathbf{P}_J W) - w| \leq \varepsilon_0^{-3/2}\varepsilon\right) + \sqrt{p_0}.$$

Let us fix a vector w_0 and a number w. We can rewrite

$$(3.11) \qquad\qquad (w_0, P_J W) = \sum_{i \in J}(a_i x_i + b_i y_i),$$

where $||a||_2^2 + ||b||_2^2 = 1$. From Lemma A.4 and Remark A.5 we know that at least $[2c_0 n]$ coordinates of vector $a \in Incomp(\delta, \tau)$ satisfy

$$\frac{\tau}{\sqrt{2n}} \leq |a_k| \leq \frac{1}{\sqrt{\delta n}}.$$

We denote the set of coordinates of a with this property by spread(a). By construction of J we can conclude that $|\text{spread}(a)| = [c_0 n]$. By Lemma A.6 we can reduce our sum (3.11) to the set spread(a). Now we will find the properties of $|b_i|$. We can decompose the set spread(a) into two sets spread$(a) = I_1 \cup I_2$:
a) $I_1 = \{i \in \text{spread}(a) : |b_i|\sqrt{n} \to \infty \text{ as } n \to \infty\}$;
c) $I_2 = \{i \in \text{spread}(a) : |b_i| = O(n^{-1/2})\}$;

From $||b||_2^2 < 1$ it follows that $|I_1| = o(n)$. For $i \in I_2$ we have $|a_i^{-1} b_i| = O(1)$.
By Lemma A.6 we have

$$\mathbb{P}(|\sum_{i \in \mathrm{spread}(a)} (a_i x_i + b_i y_i) - w| < 2\varepsilon_0^{-3/2}\varepsilon) \le \mathbb{P}(|\sum_{i \in I_2}(a_i x_i + b_i y_i) - w'| < 2\varepsilon_0^{-3/2}\varepsilon).$$

We can apply Statement 3.4

$$\mathbb{P}(|\sum_{i \in I_2}(a_i x_i + b_i y_i) - w'| < 2\varepsilon_0^{-3/2}\varepsilon) \le \frac{C_1 \varepsilon_0^{-3/2}\varepsilon}{(1-\rho^2)^{1/2}} + C_2(1-\rho^2)^{-3/2}n^{-1/2}.$$

It follows that

$$\mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap ||\mathbf{B}|| \le 3K\sqrt{n}) \le$$

$$\le \left(\frac{C_1 \varepsilon_0^{-3/2}\varepsilon}{(1-\rho^2)^{1/2}} + C_2(1-\rho^2)^{-3/2}n^{-1/2} \right)^{1/2} + p_0^{1/4} + 2\sqrt{p_0}.$$

We take $\varepsilon_0 = \varepsilon^{1/2}$ and conclude that

$$\mathbb{P}_{\mathbf{B},W}(\Omega_0 \cap ||\mathbf{B}|| \le 3K\sqrt{n}) \le C(\rho)\varepsilon^{1/8} + C'(\rho)n^{-1/8},$$

where $C(\rho), C'(\rho)$ are some constants which depend on ρ, K and M_4.   □

Proof. (proof of Theorem 3.1) The result of the theorem follows from Lemmas 3.7, 3.8 and 3.9.   □

Remark 3.14. *It not very difficult to show that we can change matrix $z\mathbf{I}$ in Theorem 3.1 by arbitrary non-random matrix $\mathbf{M}$ with $||\mathbf{M}|| \le K\sqrt{n}$. We can also assume that $\mathbb{E} X_{ij}^2 \ge 1$. Results of section 3.2 are based on Lemmas A.8 and A.9 which doesn't depend on shifts. It is easy to see that Statement 3.13 still holds true if we assume that $\varepsilon < n^{-Q}$ for some $Q > 0$. Then we can reformulate Theorem 3.1 in the following way: there exist some constants $A, B > 0$ such that*

$$\mathbb{P}(s_n(\mathbf{X} + \mathbf{M}) \le \varepsilon n^{-A}, ||\mathbf{X} + \mathbf{M}|| \le K\sqrt{n}) \le C(\rho)n^{-B}.$$

4. Uniform integrability of logarithm

In this section we prove the next result

Theorem 4.1. *Under the condition ($\mathbf{C0}$) $\log(\cdot)$ is uniformly integrable in probability with respect to $\{\nu_n\}_{n \ge 1}$.*

Before we need several lemmas about the behavior of the singular values

Lemma 4.2. *If condition ($\mathbf{C0}$) holds then there exists constant $K := K(\rho)$ such that $\mathbb{P}(s_1(\mathbf{X}) \ge K\sqrt{n}) = o(1)$.*

Proof. We can decompose matrix $\mathbf{X}$ into symmetric and skew-symmetric matrices:

$$\mathbf{X} = \frac{\mathbf{X} + \mathbf{X}^T}{2} + \frac{\mathbf{X} - \mathbf{X}^T}{2} = \mathbf{X}_1 + \mathbf{X}_2.$$

In [15, Theorem 2.3.23] it is proved that for some $K_1 > \sqrt{2(1 + \rho)}$

(4.1)
$$\mathbb{P}(s_1(\mathbf{X}_1) \geq K_1\sqrt{n}) = o(1).$$

and for some $K_2 > \sqrt{2(1 - \rho)}$

(4.2)
$$\mathbb{P}(s_1(\mathbf{X}_2) \geq K_2\sqrt{n}) = o(1)$$

Set $K = 2\max(K_1, K_2)$. From (4.1), (4.2) and inequality

$$s_1(\mathbf{X}) \leq s_1(\mathbf{X}_1) + s_1(\mathbf{X}_2)$$

it follows that

$$\mathbb{P}(s_1(\mathbf{X}) \geq K\sqrt{n}) \leq \mathbb{P}\left(\left\{s_1(\mathbf{X}_1) \geq \frac{K\sqrt{n}}{2}\right\} \cup \left\{s_1(\mathbf{X}_2) \geq \frac{K\sqrt{n}}{2}\right\}\right)$$

$$\leq \mathbb{P}\left(s_1(\mathbf{X}_1) \geq \frac{K\sqrt{n}}{2}\right) + \mathbb{P}\left(s_1(\mathbf{X}_2) \geq \frac{K\sqrt{n}}{2}\right) = o(1).$$

$\square$

Remark 4.3. *Suppose that elements of $\mathbf{X}_n$ depend on n, but satisfy conditions* (**C0**) *and* $|x_{ij}| \leq \delta_n\sqrt{n}, \mathbb{E}\, x_{ij}^2 \leq 1$ *and* $\mathbb{E}\,|x_{ij}|^l \leq b(\delta_n\sqrt{n})^{l-1}$ *for some $b > 0, l \geq 3$ and $\delta_n \to 0$ with the convergence rate slower that any preassigned one as $n \to \infty$. Then for some $K > 0$ it can be shown that*

$$\mathbb{P}(s_1(\mathbf{X}) \geq K\sqrt{n}) = o(n^{-l}).$$

For the proof see [2, Theorem 5.1].

Lemma 4.4. *If condition* (**C0**) *holds then there exist $c > 0$ and $0 < \gamma < 1$ such that a.s. for $n \gg 1$ and $n^{1-\gamma} \leq i \leq n - 1$*

$$s_{n-i}(n^{-1/2}\mathbf{X} - z\mathbf{I}) \geq c\frac{i}{n}.$$

Proof. Set $s_i := s_i(n^{-1/2}\mathbf{X} - z\mathbf{I})$. Up to increasing γ, it is sufficient to prove the statement for all $2(n-1)^{1-\gamma} \leq i \leq n - 1$ for some $\gamma \in (0, 1)$ to be chosen later. We fix some $2(n-1)^{1-\gamma} \leq i \leq n - 1$ and consider the matrix $\mathbf{A}'$ formed by the first $m := n - \lceil i/2 \rceil$ rows of $\sqrt{n}\mathbf{A}$. Let $s_1' \geq ... \geq s_m'$ be the singular values of $\mathbf{A}'$. We get

$$n^{-1/2}s_{n-i}' \leq s_{n-i}.$$

By R_i we denote the row of $\mathbf{A}'$ and $H_i = \text{span}(R_j, j = 1, ..., m, j \neq i)$. By Lemma A.10 we obtain

$$s_1'^{-2} + ... + s_{n-\lceil i/2 \rceil}'^{-2} = \text{dist}_1^{-2} + ... + \text{dist}_{n-\lceil i/2 \rceil}^{-2}.$$

We have

(4.3)
$$\frac{i}{2n}s_{n-i}^{-2} \leq \frac{i}{2}s_{n-i}'^{-2} \leq \sum_{j=n-i}^{n-\lceil i/2 \rceil} s_j'^{-2} \leq \sum_{j=1}^{n-\lceil i/2 \rceil} \text{dist}_j^{-2},$$

where $\text{dist}_j := \text{dist}(R_j, H_j)$. To estimate $\text{dist}(R_j, H_j)$ we would like to apply Lemma A.11, but we can't do it directly, because R_j and H_j are not independent. Let's consider the case $j = 1$ only. To estimate distance dist_1 we decompose matrix $\mathbf{A}'$ into the blocks

$$\mathbf{A}' = \begin{pmatrix} a_{1,1} & Y \\ X & \mathbf{B} \end{pmatrix},$$

where $X \in \mathbb{R}^{m-1}, Y^T \in \mathbb{R}^{n-1}$ and $\mathbf{B}$ is an $m - 1 \times n - 1$ matrix formed by rows $B_1, ..., B_{m-1}$. We denote by $H'_1 = \text{span}(B_1, ..., B_{m-1})$. From definition of distance

$$\text{dist}(R_1, H_1) = \inf_{v \in H_1} ||R_1 - v||_2 \geq \inf_{u \in H'} ||Y - u||_2 = \text{dist}(Y, H'_1)$$

and

$$\dim(H'_1) \leq \dim(H_1) \leq n - 1 - i/2 \leq n - 1 - (n-1)^{1-\gamma}.$$

Now vector Y and hyperplane H'_1 are independent. Fixing realization of H'_1, by Lemma A.11, with n, R, H replaced with $n - 1, Y, H'_1$ respectively, we can obtain that

$$\mathbb{P}(\text{dist}(Y, H'_1) \leq \frac{1}{2}\sqrt{n - 1 - \dim(H'_1)}) \leq \exp(-(n-1)^\delta).$$

Using this inequality it is easy to show that

$$\mathbb{P}\left(\bigcup_{n \gg 1} \bigcup_{i=\lceil 2(n-1)^{1-\gamma}\rceil}^{n-1} \bigcup_{j=1}^{n-\lceil i/2\rceil} \left\{ \text{dist}\left(R_j, H_j\right) \leq \frac{1}{2}\sqrt{\frac{i}{2}} \right\} \right) < \infty.$$

Now by Borel-Cantelli lemma and (4.3) we can conclude the statement of the lemma. $\square$

Remark 4.5. *Lemma 4.4 holds true if we assume that* $\mathbb{E}\,X_{ij} \neq 0$ *and* $\mathbb{E}\,X_{ij}^2 = 1 + o(1)$.

Proof. (Proof of Theorem 4.1) To prove Theorem 4.1 we need to show that there exist $p, q > 0$ such that

$$(4.4) \qquad\qquad \lim_{t\to\infty} \overline{\lim_{n\to\infty}} \,\mathbb{P}\left(\int_0^\infty x^p \nu_n(dx) > t \right) = 0$$

and

$$(4.5) \qquad\qquad \lim_{t\to\infty} \overline{\lim_{n\to\infty}} \,\mathbb{P}\left(\int_0^\infty x^{-q} \nu_n(dx) > t \right) = 0.$$

By Lemma 4.2 there exists set $\Omega_0 := \Omega_{0,n} = \{\omega \in \Omega : s_1(\mathbf{X}) \leq Kn^{1/2}\}$ such that

$$(4.6) \qquad\qquad \mathbb{P}(\Omega_0) = 1 - o(1).$$

We conclude (4.4) from (4.6) for $p = 2$.

We denote $\Omega_1 := \Omega_{1,n} = \{\omega \in \Omega : s_{n-i} > c\frac{i}{n}, n^{1-\gamma} \le i \le n-1\}$. Let us consider the set $\Omega_2 := \Omega_{2,n} = \Omega_1 \cap \{\omega : s_n \ge n^{-B-1/2}\}$, where $B > 0$. We decompose probability from (4.5) into two terms

$$\mathbb{P}\left(\int_0^\infty x^{-q}\nu_n(dx) > t\right) = \mathbb{I}_1 + \mathbb{I}_2,$$

where

$$\mathbb{I}_1 := \mathbb{P}\left(\int_0^\infty x^{-q}\nu_n(dx) > t, \Omega_2\right),$$

$$\mathbb{I}_2 := \mathbb{P}\left(\int_0^\infty x^{-q}\nu_n(dx) > t, \Omega_2^c\right).$$

We can estimate $\mathbb{I}_2$ by

$$\mathbb{I}_2 \le \mathbb{P}(s_n(\mathbf{X} - \sqrt{n}z\mathbf{I}) \le n^{-A}, \Omega_0) + \mathbb{P}(\Omega_0^c) + \mathbb{P}(\Omega_1^c).$$

From Theorem 3.1 it follows that

(4.7) $$\mathbb{P}(s_n(\mathbf{X} - \sqrt{n}z\mathbf{I}) \le n^{-B}, \Omega_0) \le C(\rho)n^{-1/8}.$$

By Lemma 4.4

(4.8) $$\varlimsup_{n\to\infty} \mathbb{P}(\Omega_1^c) = 0.$$

From (4.6), (4.7) and (4.8) we conclude

$$\varlimsup_{n\to\infty} \mathbb{I}_2 = 0.$$

To prove (4.5) it remains to bound $\mathbb{I}_1$. From Markov inequality

$$\mathbb{I}_1 \le \frac{1}{t}\mathbb{E}\left[\int_0^\infty x^{-q}\nu_n(dx)\mathbb{1}(\Omega_2)\right].$$

By definition of Ω_2

$$\mathbb{E}\left[\int x^{-q}\nu_n(dx)\mathbb{1}(\Omega_2)\right] \le \frac{1}{n}\sum_{i=1}^{n-\lceil n^{1-\gamma}\rceil} s_i^{-q} + \frac{1}{n}\sum_{i=n-\lceil n^{1-\gamma}\rceil+1}^{n} s_i^{-q}$$

$$\le 2n^{q(B+1/2)-\gamma} + c^{-q}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{n}{i}\right)^q \le 2n^{q(B+1/2)-\gamma} + c^{-q}\int_0^1 s^{-q}ds.$$

If $0 < q < \min(1, \gamma/(B+1/2))$ then the last integral is finite. $\square$

5. Convergence of singular values

Let function $\mathcal{F}_n(x,z)$ be an empirical distribution function of singular values $s_1 \ge ... \ge s_n$ of matrix $n^{-1/2}\mathbf{X} - z\mathbf{I}$ which corresponds to measure $\nu_n(z, \cdot)$.

Let us recall definition of Stieltjes transform

Definition 5.1. *The Stieltjes transform of measure $m(\cdot)$ on $\mathbb{R}$ is*

$$S(\alpha) = \int_{\mathbb{R}} \frac{m(dx)}{x-\alpha}, \quad \alpha \in \mathbb{C}^+.$$

In this section we prove the following theorem

Theorem 5.2. *Assume that condition* (**C0**) *holds true. There exists non-random distribution function* $\mathcal{F}(x, z)$ *such that for all continues and bounded functions* $f(x)$, *a.a.* $z \in \mathbb{C}$ *and all* $\varepsilon > 0$

$$\mathbb{P}\left(\left|\int_{\mathbb{R}} f(x)d\mathcal{F}_n(x, z) - \int_{\mathbb{R}} f(x)d\mathcal{F}(x, z)\right| > \varepsilon\right) \to 0 \ as \ n \to \infty,$$

Proof. First we show that family $\{\mathcal{F}(z, x)\}_{n \geq 1}$ is tight. From strong law of large numbers it follows that

$$\int_0^\infty x^2 d\mathcal{F}(x, z) \leq \frac{1}{n^2}\sum_{i,j=1}^n X_{ij}^2 \to 1 \text{ as } n \to \infty.$$

Using this and the fact that $s_i(n^{-1/2}\mathbf{X} - z\mathbf{I}) \leq s_i(n^{-1/2}\mathbf{X}) + |z|$ we conclude tightness of $\{\mathcal{F}_n(z, x)\}_{n \geq 1}$. If we show that $\mathcal{F}_n$ weakly converges in probability to some function $\mathcal{F}$, then $\mathcal{F}$ will be distribution function.

Introduce the following $2n \times 2n$ matrices

$$(5.1)\qquad \mathbf{V} = \begin{pmatrix} \mathbf{O}_n & n^{-1/2}\mathbf{X} \\ n^{-1/2}\mathbf{X}^T & \mathbf{O}_n \end{pmatrix}, \quad \mathbf{J}(z) = \begin{pmatrix} \mathbf{O}_n & z\mathbf{I} \\ \overline{z}\mathbf{I} & \mathbf{O}_n \end{pmatrix}$$

where $\mathbf{O}_n$ denotes $n \times n$ matrix with zero entries. Consider matrix

$$\mathbf{V}(z) := \mathbf{V} - \mathbf{J}(z).$$

It is known that eigenvalues of $\mathbf{V}(z)$ are singular values of $n^{-1/2}\mathbf{X} - z\mathbf{I}$ with signs $\pm$.

It is easy to see that empirical distribution function $F_n(x, z)$ of eigenvalues of matrix $\mathbf{V}(z)$ can be written in the following way

$$F_n(x, z) = \frac{1}{2n}\sum_{i=1}^n \mathbb{1}\{s_i \leq x\} + \frac{1}{2n}\sum_{i=1}^n \mathbb{1}\{-s_i \leq x\}.$$

There is one to one correspondence between $\mathcal{F}_n(x, z)$ and $F_n(x, z)$

$$F_n(x, z) = \frac{1 + \mathrm{sgn}(x)\mathcal{F}_n(|x|, z)}{2}$$

So it is enough to show that there exists non-random distribution function $F(x, z)$ such that for all continues and bounded functions $f(x)$, and a.a. $z \in \mathbb{C}$

$$(5.2)\qquad \mathbb{P}\left(\left|\int_{\mathbb{R}} f(x)dF_n(x, z) - \int_{\mathbb{R}} f(x)dF(x, z)\right| > \varepsilon\right) \to 0 \text{ as } n \to \infty.$$

We denote Stieltjes transforms of F_n and F by $S_n(x, z)$ and $S(x, z)$ respectively. Due to the relations between distribution functions and Stieltjes transforms, (5.2) will follow from

$$(5.3)\qquad \mathbb{P}(|S_n(\alpha, z) - S(\alpha, z)| > \varepsilon) \to 0 \text{ as } n \to \infty,$$

for a.a. $z \in \mathbb{C}$ and all $\alpha \in \mathbb{C}^+$.

Set

(5.4) $$\mathbf{R}(\alpha, z) := (\mathbf{V}(z) - \alpha \mathbf{I}_{2n})^{-1}.$$

By definition $S_n(\alpha, z) = \frac{1}{2n} \operatorname{Tr} \mathbf{R}(\alpha, z)$. We introduce the following function

$$s_n(\alpha, z) := \mathbb{E} \, S_n(\alpha, z) = \frac{1}{2n} \sum_{i=1}^{2n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{ii},$$

One can show that

$$s_n(\alpha, z) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{ii} = \frac{1}{n} \sum_{i=n+1}^{2n} \mathbb{E}[\mathbf{R}(\alpha, z)]_{ii}$$

By Chebyshev inequality and Lemma A.12 it is staighforward to check that

(5.5) $$|s_n(\alpha, z) - s(\alpha, z)| \to 0 \text{ as } n \to \infty.$$

implies (5.3).

By resolvent equality we may write

$$1 + \alpha s_n(\alpha, z) = \frac{1}{2n} \mathbb{E} \operatorname{Tr}(\mathbf{V} \mathbf{R}(\alpha, z)) - z t_n(\alpha, z) - \bar{z} u_n(\alpha, z).$$

Introduce the notation

$$\mathbb{A} := \frac{1}{2n} \mathbb{E} \operatorname{Tr}(\mathbf{V} \mathbf{R})$$

and represent $\mathbb{A}$ as follows

$$\mathbb{A} = \frac{1}{2} \mathbb{A}_1 + \frac{1}{2} \mathbb{A}_2,$$

where

$$\mathbb{A}_1 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\mathbf{V} \mathbf{R}]_{ii}, \quad \mathbb{A}_2 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\mathbf{V} \mathbf{R}]_{i+n, i+n}.$$

First we consider $\mathbb{A}_1$. By definition of the matrix $\mathbf{V}$, we have

$$\mathbb{A}_1 = \frac{1}{n^{3/2}} \sum_{j,k=1}^{n} \mathbb{E} \, X_{jk} R_{k+n, j}.$$

Note that

$$\frac{\partial \mathbf{R}}{\partial X_{jk}} = -\frac{1}{\sqrt{n}} \mathbf{R}[e_j e_{k+n}^T] \mathbf{R}.$$

Applying Lemma A.14 we obtain

$$\mathbb{A}_1 = \mathbb{B}_1 + \mathbb{B}_2 + \mathbb{B}_3 + \mathbb{B}_4 + r_n(\alpha, z).$$

where

$$\mathbb{B}_1 = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}]_{k+n,j} = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}(R_{k+n,j})^2$$

$$\mathbb{B}_2 = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{k+n} e_j^T]\mathbf{R}]_{k+n,j} = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{jj} R_{k+n,k+n}$$

$$\mathbb{B}_3 = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_k e_{j+n}^T]\mathbf{R}]_{k+n,j} = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{k+n,k} R_{j+n,j}$$

$$\mathbb{B}_4 = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{j+n} e_k^T]\mathbf{R}]_{k+n,j} = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{kj} R_{k+n,j+n}.$$

Without loss of generality we can assume further that $\mathbb{E}\, X_{11}^2 = 1$ because the impact of diagonal is of order $O(n^{-1})$.

From $||\mathbf{R}||_{HS} \leq \sqrt{n}||\mathbf{R}|| \leq \sqrt{n}v^{-1}$ it follows

$$|\mathbb{B}_1| \leq \frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, X_{jk}^2\, \mathbb{E}(R_{k+n,j})^2 \leq \frac{1}{nv^2}.$$

Similarly

$$|\mathbb{B}_4| \leq \frac{1}{v^2 n}.$$

By Lemma A.12 $\mathbb{B}_2 = -s_n^2(\alpha, z) + \varepsilon(\alpha, z)$. By Lemma A.13 $\mathbb{B}_3 = -\rho t_n^2(\alpha, z) + \varepsilon(\alpha, z)$. We obtain that

$$\mathbb{A}_1 = -s_n^2(\alpha, z) - \rho t_n^2(\alpha, z) + \delta_n(\alpha, z).$$

Now we consider the term $\mathbb{A}_2$. By definition of the matrix $\mathbf{V}$, we have

$$\mathbb{A}_2 = \frac{1}{n^{3/2}} \sum_{j,k=1}^{n} \mathbb{E}\, X_{jk} R_{j,k+n}.$$

By Lemma A.14 we obtain that

$$(5.6) \qquad \mathbb{A}_2 = \mathbb{C}_1 + \mathbb{C}_2 + \mathbb{C}_3 + \mathbb{C}_4 + r_n(\alpha, z).$$

where

$$\mathbb{C}_1 = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}]_{j,k+n} = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{jj} R_{k+n,k+n}$$

$$\mathbb{C}_2 = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{k+n} e_{j}^T]\mathbf{R}]_{j,k+n} = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}(R_{j,k+n})^2$$

$$\mathbb{C}_3 = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_k e_{j+n}^T]\mathbf{R}]_{j,k+n} = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{jk} R_{j+n,k+n}$$

$$\mathbb{C}_4 = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{j+n} e_{k}^T]\mathbf{R}]_{j,k+n} = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{j,j+n} R_{k,k+n}.$$

It is easy to show that

$$|\mathbb{C}_2| \le \frac{1}{v^2 n}, \quad |\mathbb{C}_3| \le \frac{1}{v^2 n}.$$

By Lemma A.12 $\mathbb{C}_1 = -s_n^2(\alpha, z) + \varepsilon_n(\alpha, z)$. By Lemma A.13 $\mathbb{C}_4 = -\rho u_n^2(\alpha, z) + \varepsilon_n(\alpha, z)$. We obtain that

$$\mathbb{A}_2 = -s_n^2(\alpha, z) - \rho u_n^2(\alpha, z) + \delta_n(\alpha, z).$$

So we have that

$$\mathbb{A} = -s_n^2(\alpha, z) - \frac{\rho}{2} t_n^2(\alpha, z) - \frac{\rho}{2} u_n^2(\alpha, z) + \varepsilon_n(\alpha, z).$$

No we will investigate the term $z t_n(\alpha, z)$ which we may represent as follows

$$\alpha t_n(\alpha, z) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[\mathbf{V}(z)\mathbf{R}]_{j+n,j} = \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[\mathbf{V}\mathbf{R}]_{j+n,j} - \bar{z} s_n(\alpha, z).$$

By definition of the matrix $\mathbf{V}$, we have

$$\alpha t_n(\alpha, z) = \frac{1}{n^{3/2}} \sum_{j,k=1}^{n} \mathbb{E}\, X_{jk} R_{j,k} - \bar{z} s_n(\alpha, z) =$$

$$\mathbb{D}_1 + \mathbb{D}_2 + \mathbb{D}_3 + \mathbb{D}_4 - \bar{z} s_n(\alpha, z) + r_n(\alpha, z),$$

where

$$\mathbb{D}_1 = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_j e_{k+n}^T]\mathbf{R}]_{j,k} = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{j,j} R_{k+n,k}$$

$$\mathbb{D}_2 = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{k+n} e_{j}^T]\mathbf{R}]_{j,k} = -\frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{j,k+n} R_{j,k}$$

$$\mathbb{D}_3 = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_k e_{j+n}^T]\mathbf{R}]_{j,k} = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{j,k} R_{j+n,k}$$

$$\mathbb{D}_4 = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}[\mathbf{R}[e_{j+n} e_{k}^T]\mathbf{R}]_{j,k} = -\frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E}\, R_{j,j+n} R_{k,k}.$$

By similar arguments as before we can prove that

$$|\mathbb{D}_2| \leq \frac{1}{v^2 n}, \quad |\mathbb{D}_3| \leq \frac{1}{v^2 n}$$

and $\mathbb{D}_1 = -s_n(\alpha, z)t_n(\alpha, z) + \varepsilon_n(\alpha, z)$, $\mathbb{D}_4 = -\rho s_n(\alpha, z)u_n(\alpha, z) + \varepsilon_n(\alpha, z)$. We obtain that

$$\alpha t_n(\alpha, z) = -s_n(\alpha, z)t_n(\alpha, z) - \rho s_n(\alpha, z)u_n(\alpha, z) - \overline{z}s_n(\alpha, z) + \delta_n(\alpha, z).$$

Similar we can prove that

$$\alpha u_n(\alpha, z) = -s_n(\alpha, z)u_n(\alpha, z) - \rho s_n(\alpha, z)t_n(\alpha, z) - z s_n(\alpha, z) + \delta_n(\alpha, z).$$

So we have the system of equations

(5.7) $$1 + \alpha s_n(\alpha, z) + s_n^2(\alpha, z) =$$
$$= -\frac{\rho}{2}t_n^2(\alpha, z) - \frac{z}{2}t_n(\alpha, z) - \frac{\rho}{2}u_n^2(\alpha, z) - \frac{\overline{z}}{2}u_n(\alpha, z) + \delta_n(\alpha, z)$$

(5.8) $$\alpha t_n(\alpha, z) =$$
$$= -s_n(\alpha, z)t_n(\alpha, z) - \rho s_n(\alpha, z)u_n(\alpha, z) - \overline{z}s_n(\alpha, z) + \delta_n(\alpha, z)$$

(5.9) $$\alpha u_n(\alpha, z) =$$
$$= -s_n(\alpha, z)u_n(\alpha, z) - \rho s_n(\alpha, z)t_n(\alpha, z) - z s_n(\alpha, z) + \delta_n(\alpha, z).$$

It follows from (5.8) and (5.9) that

$$(\alpha + s_n)(zt_n + \rho t_n^2) = -s_n(z\rho u_n + \overline{z}\rho t) - \rho^2 s_n t_n u_n - |z|^2 s_n + \delta_n(\alpha, z)$$
$$(\alpha + s_n)(\overline{z}u_n + \rho u_n^2) = -s_n(z\rho u_n + \overline{z}\rho t) - \rho^2 s_n t_n u_n - |z|^2 s_n + \delta_n(\alpha, z).$$

So, we can rewrite (5.7)

(5.10) $$1 + \alpha s_n(\alpha, z) + s_n^2(\alpha, z) + \rho^2 t_n^2(\alpha, z) + z t_n(\alpha, z) = \delta_n(\alpha, z).$$

From equations (5.8) and (5.9) we can write equation for t_n

(5.11) $$\left(\alpha + s_n - \frac{|\rho|^2 s_n^2}{\alpha + s_n}\right) t_n = \frac{\rho z s_n^2}{\alpha + s_n} - \overline{z}s_n + \delta_n(\alpha, z).$$

We denote

$$\Delta = \left(\alpha + s_n - \frac{|\rho|^2 s_n^2}{\alpha + s_n}\right).$$

After simple calculations we will have

$$(\alpha + s_n)(zt_n + \rho t_n^2) =$$
$$- s_n\left(\frac{2\rho^2|z|^2 s_n^2}{(\alpha + s_n)\Delta} - \frac{\overline{z}^2\rho s_n}{\Delta} - \frac{z^2\overline{\rho}s_n}{\Delta}\right)$$
$$- |\rho|^2 s_n\left(\frac{\rho z s_n^2}{(\alpha + s)\Delta} - \frac{\overline{z}s_n}{\Delta}\right)\left(\frac{\rho\overline{z}s_n^2}{(\alpha + s)\Delta} - \frac{z s_n}{\Delta}\right) - |z|^2 s_n + \delta_n(\alpha, z).$$
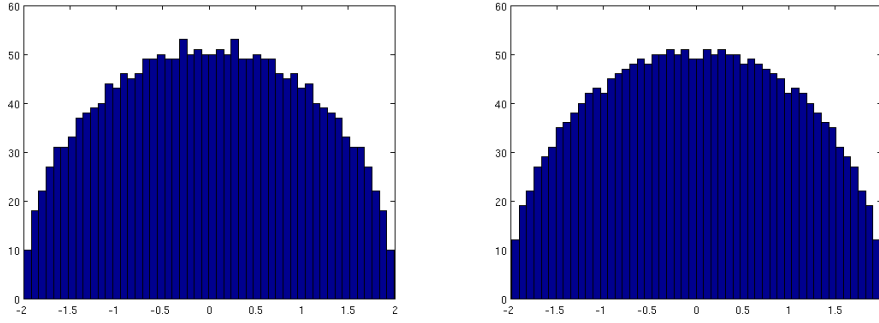
FIGURE 4. Histogram of eigenvalues of matrix $\mathbf{V}$ for $n = 1000$. entries are Gaussian random variables. On the left $\rho = 0$ (Circular law case). On the right $\rho = 0.5$ (Elliptic law case).

We denote $y_n := s_n$ and $w_n := \alpha + (\rho t_n^2 + z t_n)/y_n$. We can rewrite equations (5.7), (5.8) and (5.9)

$$(5.12) \qquad 1 + w_n y_n + y_n^2 = \delta_n(\alpha, z)$$

$$(5.13) \qquad w_n = \alpha + \frac{\rho t_n^2 + z t_n}{y_n}$$

$$(5.14) \qquad (\alpha + s_n)(z t_n + \rho t_n^2) =$$
$$- s_n \left(\frac{2\rho^2 |z|^2 y_n^2}{(\alpha + y_n)\Delta} - \frac{\bar{z}^2 \rho y_n}{\Delta} - \frac{z^2 \rho y_n}{\Delta} \right) - |z|^2 y_n$$
$$- |\rho|^2 y_n \left(\frac{\rho z y_n^2}{(\alpha + y_n)\Delta} - \frac{\bar{z} y_n}{\Delta} \right) \left(\frac{\rho \bar{z} y_n^2}{(\alpha + y_n)\Delta} - \frac{z y_n}{\Delta} \right) + \delta_n(\alpha, z).$$

Remark 5.3. *If $\rho = 0$ then we can rewrite* (5.12), (5.13), *and* (5.14)

$$1 + w_n y_n + y_n^2 = \delta_n(\alpha, z)$$
$$w_n = \alpha + \frac{z t_n}{y_n}$$
$$(w_n - \alpha) + (w_n - \alpha)^2 y_n - |z|^2 s_n = \delta_n(\alpha, z).$$

This equations determine the Circular law, see [9].

We can see that the first equation (5.12) doesn't depend on ρ. So the first equation will be the same for all models of random matrices described in the introduction. On the Figure 4 we draw the distribution of eigenvalues of matrix $\mathbf{V}$ for $\rho = 0$ (Circular law case) and $\rho = 0.5$ (Elliptic law case).

Now we prove convergence of s_n to some limit s_0. Let $\alpha = u + iv, v > 0$. Using (5.10) we write

$$\alpha(s_n - s_m) = -(s_n - s_m)(s_n + s_m) - \rho^2(t_n - t_m)(t_n + t_m) - z(t_m - t_m) + \varepsilon_{n,m}.$$

By triangle inequality and the fact that $|s_n| \leq v^{-1}$

$$(5.15) \quad |s_n - s_m| \leq \frac{2|s_n - s_m|}{v^2} + \frac{\rho^2 |t_n - t_m||t_n + t_m|}{v} + \frac{|z||t_n - t_m|}{v} + \frac{|\varepsilon_{n,m}|}{v}.$$

From (5.11) it follows that

$$((\alpha + s_n)^2 - \rho^2 s_n^2)t_n = \rho z s_n^2 - \bar{z}\alpha s_n - \bar{z} s_n^2 + \varepsilon_n.$$

We denote $\Delta_n := ((\alpha + s_n)^2 - \rho^2 s_n^2)$. By triangle inequality

$$(5.16) \qquad |\Delta_m||t_n - t_m| \leq |t_m||\Delta_n - \Delta_m|$$

$$+ \frac{2|\rho||s_n - s_m| + 2|z||s_n - s_m|}{v} + |z||\alpha||s_n - s_m| + |\varepsilon_{n,m}|.$$

We can find lower bound for $|\Delta_m|$:

$$(5.17) \qquad |\Delta_m| = |\alpha + (1 - \rho)s_m||\alpha + (1 + \rho)s_m|$$

$$\geq \operatorname{Im}(\alpha + (1 - \rho)s_m)\operatorname{Im}(\alpha + (1 + \rho)s_m) \geq v^2,$$

where we have used the fact that $\operatorname{Im} s_m \geq 0$. From definition of Δ_n it is easy to see that

$$(5.18) \qquad |\Delta_n - \Delta_m| \leq 2|\alpha||s_n - s_m| + \frac{2(1 + \rho^2)|s_n - s_m|}{v}.$$

We can take $|u| \leq C$, then $|\alpha| \leq v + C$. From (5.15),(5.16),(5.17) and (5.18) it follows that there exists constant C', which depends on ρ, C, z, such that

$$|s_n - s_m| \leq \frac{C'}{v}|s_n - s_m| + |\varepsilon'_{n,m}(\alpha, z)|.$$

We can find v_0 such that

$$\frac{C'}{v} < 1 \quad \text{for all } v \geq v_0.$$

Since $\varepsilon'_{n,m}(\alpha, z)$ converges to zero uniformly for all $v \geq v_0, |u| \leq C$ and s_n, s_m are locally bounded analytic functions in the upper half-plane we may conclude by Montel's Theorem (see [4, Theorem 2.9]) that there exists an analytic function s_0 in the upper half-plane such that $\lim s_n = s_0$. Since s_n are Nevanlinna functions, (that is analytic functions mapping the upper half-plane into itself) s_0 will be a Nevanlinna function too and there exists non-random distribution function $F(z, x)$ such that

$$s_0(\alpha) = \int \frac{dF(z, x)}{x - \alpha}.$$

The function s_0 satisfies the equations (5.12), (5.13), and (5.14). $\qquad \square$

6. Acknowledgment

Appendix A. Appendix

Theorem A.1. *(Central Limit Theorem) Let $Z_1, ..., Z_n$ be independent random variables with $\mathbb{E}\, Z_i = 0$ and finite third moment, and let $\sigma^2 = \sum_{i=1}^n \mathbb{E}\, |Z_i|^2$. Consider a standard normal variable g. The for every $t > 0$:*

$$\left| \mathbb{P}\left(\frac{1}{\sigma} \sum_{i=1}^n Z_i \leq t \right) - \mathbb{P}\left(g \leq t \right) \right| \leq C\sigma^{-3} \sum_{i=1}^n \mathbb{E}\, |Z_i|^3,$$

where C is an absolute constant.

Lemma A.2. *Let event $E(X, Y)$ depends on independent random vectors X and Y then*

$$\mathbb{P}(E(X, Y)) \leq (\mathbb{P}(E(X, Y), E(X, Y')))^{1/2},$$

where Y' is an independent copy of Y.

Proof. See in [5]. $\square$

Lemma A.3. *Let $Z_1, ..., Z_n$ be a sequence of random variables and $p_1, ..., p_n$ be non-negative real numbers such that*

$$\sum_{i=1}^n p_i = 1,$$

then for every $\varepsilon > 0$

$$\mathbb{P}(\sum_{i=1}^n p_i Z_i \leq \varepsilon) \leq 2 \sum_{i=1}^n p_i\, \mathbb{P}(Z_i \leq 2\varepsilon).$$

Proof. See in [17]. $\square$

Lemma A.4. *If $x \in Incomp(\delta, \tau)$ then at least $\frac{1}{2}\delta\tau^2 n$ coordinates x_k of x satisfy*

$$\frac{\tau}{\sqrt{2n}} \leq |x_k| \leq \frac{1}{\sqrt{\delta n}}.$$

Remark A.5. *We can fix some constant c_0 such that*

$$\frac{1}{4}\delta\tau^2 \leq c_0 \leq \frac{1}{4}.$$

Then for every vector $x \in Incomp(\delta, \tau)$ $|\operatorname{spread}(x)| = [2c_0 n]$.

Proof. See in [11]. $\square$

Lemma A.6. *Let $S_J = \sum_{i \in J} \xi_i$, where $J \subset [n]$, and $I \subset J$ then*

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|S_J - v| \leq \varepsilon) \leq \sup_{v \in \mathbb{R}} \mathbb{P}(|S_I - v| \leq \varepsilon).$$

Proof. Let us fix arbitrary v. From independence of ξ_i we conclude

$$\mathbb{P}(|S_J - v| \leq \varepsilon) \leq \mathbb{E}\,\mathbb{P}(|S_I + S_{J/I} - v| \leq \varepsilon|\{\xi_i\}_{i \in I}) \leq \sup_{u \in \mathbb{R}} \mathbb{P}(|S_I - u| \leq \varepsilon).$$

$\square$

Lemma A.7. *Let Z be a random variable with $\mathbb{E}\,Z^2 \geq 1$ and with finite fourth moment, and put $M_4^4 := \mathbb{E}(Z - \mathbb{E}\,Z)^4$. Then for every $\varepsilon \in (0,1)$ there exists $p = p(M_4, \varepsilon)$ such that*

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|Z - v| \leq \varepsilon) \leq p.$$

Proof. See in [11]. $\square$

Lemma A.8. *Let $\xi_1, ..., \xi_n$ be independent random variables with $\mathbb{E}\,\xi_i^2 \geq 1$ and $\mathbb{E}(\xi_k - \mathbb{E}\,\xi)^4 \leq M_4^4$, where M_4 is some finite number. Then for every $\varepsilon \in (0,1)$ there exists $p = p(M_4, \varepsilon) \in (0,1))$ such that the following holds: for every vector $x = (x_1, ..., x_n) \in S^{n-1}$, the sum $S = \sum_{i=1}^n x_k \xi_k$ satisfies*

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|S - v| \leq \varepsilon) \leq p.$$

Proof. See in [11]. $\square$

Lemma A.9. *Let $X = (X_1, ..., X_n)$ be a random vector in $\mathbb{R}^n$ with independent coordinates X_k.*
1. Suppose there exists numbers $\varepsilon_0 \geq 0$ and $L \geq 0$ such that

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|X_k - v| \leq \varepsilon) \leq L\varepsilon \quad \text{for all } \varepsilon \geq \varepsilon_0 \text{ and all } k.$$

Then

$$\sup_{v \in \mathbb{R}^n} \mathbb{P}(||X - v||_2 \leq \varepsilon\sqrt{n}) \leq (CL\varepsilon)^n \quad \text{for all } \varepsilon \geq \varepsilon_0,$$

where C is an absolute constant.
2. Suppose there exists numbers $\varepsilon > 0$ and $p \in (0,1)$ such that

$$\sup_{v \in \mathbb{R}} \mathbb{P}(|X_k - v| \leq \varepsilon) \leq L\varepsilon \quad \text{for all } k.$$

Then there exists numbers $\varepsilon_1 = \varepsilon_1(\varepsilon, p) > 0$ and $p_1 = p_1(\varepsilon, p) \in (0,1)$ such that

$$\sup_{v \in \mathbb{R}^n} \mathbb{P}(||X - v||_2 \leq \varepsilon_1\sqrt{n}) \leq (p_1)^n.$$

Proof. See [17, Lemma 3.4]. $\square$

Lemma A.10. *Let $1 \leq m \leq n$. If $\mathbf{A}$ has full rank, with rows $R_1, ..., R_m$ and $H = \text{span}(R_j, j \neq i)$, then*

$$\sum_{i=1}^m s_i(\mathbf{A})^{-2} = \sum_{i=1}^m \text{dist}(R_i, H_i)^{-2}.$$

Proof. See [16, Lemma A.4]. $\square$

Lemma A.11. *There exist $\gamma > 0$ and $\delta > 0$ such that for all $n \gg 1$ and $1 \le i \le n$, any deterministic vector $v \in \mathbb{C}$ and any subspace H of $\mathbb{C}^n$ with $1 \le \dim(H) \le n - n^{1-\gamma}$, we have, denoting $R := (X_1, ..., X_n) + v$,*

$$\mathbb{P}(\mathrm{dist}(R, H) \le \frac{1}{2}\sqrt{n - \dim(H)}) \le \exp(-n^\delta).$$

Proof. See [16, Statement 5.1]. $\qquad\qquad\square$

Lemma A.12. *Under the condition* (**C0**) *for $\alpha = u + iv, v > 0$*

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n R_{ii}(\alpha, z) - \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n R_{ii}(\alpha, z)\right)\right|^2 \le \frac{C}{nv^2}.$$

Proof. To prove this lemma we will use Girko's method. Let $\mathbf{X}^{(j)}$ be a matrix $\mathbf{X}$ with j-th row and column removed. Define matrices $\mathbf{V}^{(j)}$ and $\mathbf{V}^{(j)}(z)$ as in (5.1) and $\mathbf{R}^{(j)}$ by (5.4). It is easy to see that

$$\mathrm{rk}(\mathbf{V}(z) - \mathbf{V}^{(j)}(z)) = \mathrm{rk}(\mathbf{V} - \mathbf{V}^{(j)}) \le 4.$$

Then

$$(A.1) \quad \frac{1}{n}|\mathrm{Tr}(\mathbf{V}(z) - \alpha\mathbf{I})^{-1} - \mathrm{Tr}(\mathbf{V}^{(j)}(z) - \alpha\mathbf{I})^{-1}| \le \frac{\mathrm{rk}(\mathbf{V}(z) - \mathbf{V}^{(j)}(z))}{nv} \le \frac{4}{nv}.$$

We introduce the family of σ-algebras $\mathcal{F}_i = \sigma\{X_{j,k}, j, k > i\}$ and conditional mathematical expectation $\mathbb{E}_i = \mathbb{E}(\cdot|\mathcal{F}_i)$ with respect to this σ-algebras. We can write

$$\frac{1}{n}\mathrm{Tr}\,\mathbf{R} - \frac{1}{n}\mathbb{E}\,\mathrm{Tr}\,\mathbf{R} = \frac{1}{n}\sum_{i=1}^n \mathbb{E}_i\,\mathrm{Tr}\,\mathbf{R} - \mathbb{E}_{i-1}\,\mathrm{Tr}\,\mathbf{R} = \sum_{i=1}^n \gamma_i.$$

The sequence $(\gamma_i, \mathcal{F}_i)_{i \ge 1}$ is a martingale difference. By (A.1)

$$(A.2) \qquad |\gamma_i| = \frac{1}{n}|\mathbb{E}_i(\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(i)}) - \mathbb{E}_{i-1}(\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(i)})| \le$$

$$(A.3) \qquad \le |\mathbb{E}_i(\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(i)})| + |\mathbb{E}_{i-1}(\mathrm{Tr}\,\mathbf{R} - \mathrm{Tr}\,\mathbf{R}^{(i)})| \le \frac{C}{vn}.$$

From Burkholder inequality for martingale difference (see [13])

$$\mathbb{E}\left|\sum_{i=1}^n \gamma_i\right|^2 \le K_2\,\mathbb{E}\left(\sum_{i=1}^n |\gamma_i|^2\right)$$

and (A.2) it follows

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n R_{ii}(\alpha, z) - \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n R_{ii}(\alpha, z)\right)\right|^2 \le K_2\,\mathbb{E}(\sum_{i=1}^n |\gamma_i|^2) \le K_2\frac{C}{nv^2}.$$

$$\square$$

Lemma A.13. *Under the condition* (**C0**) *for* $\alpha = u + iv, v > 0$

$$\mathbb{E} \left| \frac{1}{n} \sum_{i=1}^{n} R_{i,i+n}(\alpha, z) - \mathbb{E} \left(\frac{1}{n} \sum_{i=1}^{n} R_{i,i+n}(\alpha, z) \right) \right|^2 \le \frac{C}{nv^4}.$$

Proof. As in Lemma A.12 we introduce matrices $\mathbf{V}^{(j)}$ and $\mathbf{R}^{(j)}$. We have

$$\mathbf{V} = \mathbf{V}^{(j)} + e_j e_j^T \mathbf{V} + \mathbf{V} e_j e_j^T + e_{j+n} e_{j+n}^T \mathbf{V} + \mathbf{V} e_{j+n} e_{j+n}^T$$

By resolvent equality $\mathbf{R} - \mathbf{R}^{(j)} = -\mathbf{R}^{(j)} (\mathbf{V}(z) - \mathbf{V}^{(j)}(z)) \mathbf{R}$

$$\frac{1}{n} \sum_{k=1}^{n} (\mathbf{R}_{k,k+n} - \mathbf{R}^{(j)}_{k,k+n}) =$$

$$= \frac{1}{n} \sum_{k=1}^{n} [\mathbf{R}^{(j)} (e_j e_j^T \mathbf{V} + e_{j+n} e_{j+n}^T \mathbf{V} + \mathbf{V} e_j e_j^T + \mathbf{V} e_{j+n} e_{j+n}^T) \mathbf{R}]_{k,k+n} =$$

$$= \mathbb{T}_1 + \mathbb{T}_2 + \mathbb{T}_3 + \mathbb{T}_4.$$

Let us consider the first term. The arguments for other terms are similar.

$$\sum_{k=1}^{n} [\mathbf{R}^{(j)} e_j e_j^T \mathbf{V} \mathbf{R}]_{k,k+n} = \operatorname{Tr} \mathbf{R}^{(j)} e_j e_j^T \mathbf{V} \mathbf{R} \mathbf{E} = \sum_{i=1}^{2n} [\mathbf{R} \mathbf{E} \mathbf{R}^{(j)}]_{ij} [e_j e_j^T \mathbf{V}]_{ji},$$

where

$$\mathbf{E} = \begin{pmatrix} \mathbf{O}_n & \mathbf{O}_n \\ \mathbf{I} & \mathbf{O}_n \end{pmatrix}.$$

From $\max(\|\mathbf{R}^{(j)}\|, \|\mathbf{R}\|) \le v^{-1}$ and Hölder inequality it follows that

$$\mathbb{E} \left| \sum_{k=1}^{n} [\mathbf{R}^{(j)} e_j e_j^T \mathbf{V} \mathbf{J} \mathbf{R}]_{k,k+n} \right|^2 \le \frac{C}{v^4}.$$

By similar arguments as in Lemma A.12 we can conclude the statement of the Lemma. $\qquad \square$

Lemma A.14. *Under the condition* (**C0**) *for* $\alpha = u + iv, v > 0$

$$\frac{1}{n^{3/2}} \sum_{j,k=1}^{n} \mathbb{E} X_{jk} R_{k+n,j} =$$

$$= \frac{1}{n^2} \sum_{j,k=1}^{n} \mathbb{E} \left[\frac{\partial \mathbf{R}}{\partial X_{jk}} \right]_{k+n,j} + \frac{\rho}{n^2} \sum_{j,k=1}^{n} \mathbb{E} \left[\frac{\partial \mathbf{R}}{\partial X_{kj}} \right]_{k+n,j} + r_n(\alpha, z),$$

where

$$|r_n(\alpha, z)| \le \frac{C}{\sqrt{n} v^3}$$

Proof. By Taylor's formula
(A.4) $\quad \mathbb{E} X f(X, Y) = f(0,0) \mathbb{E} X + f_x'(0,0) \mathbb{E} X^2 + f_y'(0,0) \mathbb{E} XY +$

$$+ \mathbb{E}(1 - \theta) [X^3 f_{xx}''(\theta X, \theta Y) + 2 X^2 Y f_{xy}''(\theta X, \theta Y) + X Y^2 f_{yy}''(\theta X, \theta Y)]$$

and

$$(A.5) \qquad \mathbb{E}\, f'_x(X, Y) = f'_x(0, 0) + \mathbb{E}(1 - \theta)[X f''_{xx}(\theta X, \theta Y) + Y f''_{xy}(\theta X, \theta Y)]$$
$$\mathbb{E}\, f'_y(X, Y) = f'_y(0, 0) + \mathbb{E}(1 - \theta)[X f''_{xy}(\theta X, \theta Y) + Y f''_{yy}(\theta X, \theta Y)],$$

where θ has uniform distribution on $[0, 1]$. From (A.4) and (A.5) for $j \neq k$

$$\left| \mathbb{E}\, X_{jk} R_{k+n,j} - \mathbb{E} \left[\frac{\partial \mathbf{R}}{\partial X_{jk}} \right]_{k+n,j} - \rho\, \mathbb{E} \left[\frac{\partial \mathbf{R}}{\partial X_{kj}} \right]_{k+n,j} \right| \leq$$

$$(|X_{jk}|^3 + |X_{jk}|) \left| \left[\frac{\partial^2 \mathbf{R}}{\partial X_{jk}^2}(\theta X_{jk}, \theta X_{kj}) \right]_{k+n,j} \right| +$$

$$(|X_{kj}|^2 |X_{jk}| + |X_{kj}|) \left| \left[\frac{\partial^2 \mathbf{R}}{\partial X_{kj}^2}(\theta X_{jk}, \theta X_{kj}) \right]_{k+n,j} \right| +$$

$$(2|X_{jk}|^2 |X_{kj}| + |X_{jk}| + |X_{kj}|) \left| \left[\frac{\partial^2 \mathbf{R}}{\partial X_{jk} \partial X_{kj}}(\theta X_{jk}, \theta X_{kj}) \right]_{k+n,j} \right|.$$

Let us consider the first term in the sum. The bounds for the second and third terms can be obtained by similar arguments. We have

$$\frac{\partial^2 \mathbf{R}}{\partial X_{jk}^2} = \frac{1}{n} \mathbf{R}(e_j e_{n+k}^T + e_{n+k} e_j^T)\mathbf{R}(e_j e_{n+k}^T + e_{n+k} e_j^T)\mathbf{R} \quad = \mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3 + \mathbb{P}_4,$$

where

$$\mathbb{P}_1 = \frac{1}{n} \mathbf{R} e_j e_{n+k}^T \mathbf{R} e_j e_{n+k}^T \mathbf{R}$$

$$\mathbb{P}_2 = \frac{1}{n} \mathbf{R} e_j e_{n+k}^T \mathbf{R} e_{n+k} e_j^T \mathbf{R}$$

$$\mathbb{P}_3 = \frac{1}{n} \mathbf{R} e_{n+k} e_j^T \mathbf{R} e_j e_{n+k}^T \mathbf{R}$$

$$\mathbb{P}_4 = \frac{1}{n} \mathbf{R} e_{n+k} e_j^T \mathbf{R} e_{n+k} e_j^T \mathbf{R}.$$

From $|\mathbf{R}_{i,j}| \leq v^{-1}$ it follows that

$$\frac{1}{n^{5/2}} \sum_{j,k=1}^{n} \mathbb{E}\, |X_{jk}|^\alpha |[\mathbb{P}_i]_{n+k,j}| \leq \frac{C}{\sqrt{n} v^3}$$

for $\alpha = 1, 3$ and $i = 1, ..., 4$. For $j = k$

$$\frac{1}{n^2} \sum_{j=1}^{n} \mathbb{E} \left[\frac{\partial \mathbf{R}}{\partial X_{jj}} \right]_{j+n,j} = \frac{1}{n^2} \sum_{j=1}^{n} (\mathbb{E}\, R_{j+n,j}^2 + \mathbb{E}\, R_{j,j} R_{j+n,j+n}) \leq \frac{C}{n v^2}.$$

So we can add this term to the sum

$$\frac{\rho}{n^2} \sum_{\substack{j,k=1 \\ j \neq k}}^{n} \mathbb{E} \left[\frac{\partial \mathbf{R}}{\partial X_{kj}} \right]_{k+n,j}.$$

$\square$

References

[1] G. Akemann, J. Baik, and P. Di Francesco. *The Oxford Handbook of Random Matrix Theory*. Oxford Unversity Press, London, 2011.

[2] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer, New York, second edition, 2010.

[3] Charles Bordenave and Djalil Chafaï. Around the circular law. *arXiv:1109.3343*.

[4] John B. Conway. *Functions of one complex variable*, volume 11. Springer-Verlag, New York, second edition, 1978.

[5] Kevin Costello. Bilinear and quadratic variants on the Littlewood-Offord problem. *Submitted*.

[6] Yan V. Fyodorov, Boris A. Khoruzhenko, and Hans-Juergen. Sommers. Universality in the random matrix spectra in the regime of weak non-hermiticity. *Ann. Inst. Henri Poincare: Phys. Theor.*, 68(4):449–489, 1998.

[7] V. L. Girko. The elliptic law. *Teor. Veroyatnost. i Primenen.*, 30(4):640–651, 1985.

[8] V. L. Girko. The strong elliptic law. Twenty years later. *Random Oper. and Stoch. Equ.*, 14(1):59–102, 2006.

[9] Friedrich Götze and Alexander Tikhomirov. The circular law for random matrices. *Ann. Probab.*, 38(4):1444–1491, 2010.

[10] Michel Ledoux. Complex hermitian polynomials: from the semi-circular law to the circular law. *Commun. Stoch. Anal.*, 2(1):27–32, 2008.

[11] Mark Rudelson and Roman Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218(2):600–633, 2008.

[12] E. B. Saff and V. Totik. *Logarithmic potentials with external fields*, volume 316. Springer-Verlag, Berlin, 1997.

[13] Albert N. Shiryaev. *Probability*, volume 95. Springer-Verlag, New York, second edition, 1996.

[14] Hans-Juergen. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein. Spectrum of large random asymmetric matrices. *Phys. Rev. Lett.*, 60:1895–1898, May 1988.

[15] Terence Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012.

[16] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.

[17] Roman Vershynin. Invertibility of symmetric random matrices. *arXiv:1102.0300*.

A. Naumov, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia

E-mail address: `naumovne@gmail.com, anaumov@math.uni-bielefeld.de`